SEQUENCE LISTING

<110> Strathmann AG & Co.

<120> Virus-Vaccine

<130> P057760

<140>
<141>

<150> 199 07 485.2
<151> 1999-02-12

<160> 12

<170> PatentIn Ver. 2.1

<210> 1
<211> 9709
<212> DNA
<213> Human immunodeficiency virus

<400> 1

```
tggaagggct aatttggtcc caaaaaagac aagagatcct tgatctgtgg atctaccaca   60
cacaaggcta cttccctgat tggcagaact acacaccagg gccagggatc agatatccac  120
tgacctttgg atggtgcttc aagttagtac cagttgaacc agagcaagta gaagaggcca  180
aataaggaga gaagaacagc ttgttacacc ctatgagcca gcatgggatg gaggacccgg  240
agggagaagt attagtgtgg aagtttgaca gcctcctagc atttcgtcac atggcccgag  300
agctgcatcc ggagtactac aaagactgct gacatcgagc tttctacaag ggactttccg  360
ctggggactt tccagggagg tgtggcctgg cgggactggg gagtggcga gccctcagat  420
gctacatata agcagctgct ttttgcctgt actgggtctc tctggttaga ccagatctga  480
gcctgggagc tctctggcta actagggaac ccactgctta agcctcaata aagcttgcct  540
tgagtgctca aagtagtgtg tgcccgtctg ttgtgtgact ctggtaacta gagatccctc  600
agaccctttt agtcagtgtg aaaatctct agcagtggcg cccgaacagg gacttgaaag  660
cgaaagtaaa gccagaggag atctctcgac gcaggactcg gcttgctgaa gcgcgcacgg  720
caagaggcga ggggcggcga ctggtgagta cgccaaaaat tttgactagc ggaggctaga  780
aggagagaga tgggtgcgag agcgtcggta ttaagcgggg gagaattaga taaatgggaa  840
aaaattcggt taaggccagg gggaaagaaa caatataaac taaaacatat agtatgggca  900
agcagggagc tagaacgatt cgcagttaat cctggccttt tagagacatc agaaggctgt  960
agacaaatac tgggacagct acaaccatcc cttcagacag gatcagaaga acttagatca 1020
ttatataata caatagcagt cctctattgt gtgcatcaaa ggatagatgt aaaagacacc 1080
aaggaagcct tagataagat agaggaagag caaaacaaaa gtaagaaaaa ggcacagcaa 1140
gcagcagctg acacaggaaa caacagccag gtcagccaaa attaccctat agtgcagaac 1200
ctccagggggc aaatggtaca tcaggccata tcacctagaa ctttaaatgc atgggtaaaa 1260
gtagtagaag agaaggcttt cagcccagaa gtaatacccä tgttttcagc attatcagaa 1320
ggagccaccc cacaagattt aaataccatg ctaaacacag tggggggaca tcaagcagcc 1380
atgcaaatgt taaaagagac catcaatgag gaagctgcag aatgggatag attgcatcca 1440
gtgcatgcag ggcctattgc accaggccag atgagagaac caaggggaag tgacatagca 1500
ggaactacta gtacccttca ggaacaaata ggatggatga cacataatcc acctatccca 1560
gtaggagaaa tctataaaag atggataatc ctgggattaa ataaaatagt aagaatgtat 1620
agccctacca gcattctgga cataagacaa ggaccaaagg aacccttttag agactatgta 1680
gaccgattct ataaaactct aagagccgag caagcttcac aagaggtaaa aaattggatg 1740
acagaaacct tgttggtcca aaatgcgaac ccagattgta agactatttt aaaagcattg 1800
ggaccaggag cgacactaga agaaatgatg acagcatgtc agggagtggg gggacccggc 1860
cataaagcaa gagttttggc tgaagcaatg agccaagtaa caaatccagc taccataatg 1920
atacagaaag gcaatttag gaaccaaaga aagactgtta agtgtttcaa ttgtggcaaa 1980
gaagggcaca tagccaaaaa ttgcagggcc cctaggaaaa agggctgttg gaaatgtgga 2040
aaggaaggac accaaatgaa agattgtact gagagacagg ctaattttttt agggaagatc 2100
tggccttccc acaagggaag gccagggaat tttcttcaga gcagaccaga gccaacagcc 2160
ccaccagaag agagcttcag gtttggggga gagacaacaa ctccctctca gaagcaggag 2220
ccgatagaca aggaactgta tcctttagct tccctcagat cactctttgg cagcgacccc 2280
tcgtcacaat aaagataggg gggcaattaa aggaagctct attagataca ggagcagatg 2340
atacagtatt agaagaaatg aatttgccag gaagatggaa accaaaaatg ataggggaa 2400
ttggaggttt tatcaaagta ggacagtatg atcagatact catagaaatc tgcggacata 2460
aagctatagg tacagtatta gtaggaccta cacctgtcaa cataattgga agaaatctgt 2520
tgactcagat tggctgcact ttaaattttc ccattagtcc tattgagact gtaccagtaa 2580
aattaaagcc aggaatggat ggcccaaaag ttaaacaatg gccattgaca gaagaaaaaa 2640
```

```
taaaagcatt agtagaaatt tgtacagaaa tggaaaagga aggaaaaatt tcaaaaattg 2700
ggcctgaaaa tccatacaat actccagtat ttgccataaa gaaaaaagac agtactaaat 2760
ggagaaaatt agtagatttc agagaactta ataagagaac tcaagatttc tgggaagttc 2820
aattaggaat accacatcct gcagggttaa aacagaaaaa atcagtaaca gtactggatg 2880
tgggcgatgc atatttttca gttcccttag ataaagactt caggaagtat actgcattta 2940
ccatacctag tataaacaat gagacaccag ggattagata tcagtacaat gtgcttccac 3000
agggatggaa aggatcacca gcaatattcc agtgtagcat gacaaaaatc ttagagcctt 3060
ttagaaaaca aaatccagac atagtcatct atcaatacat ggatgatttg tatgtaggat 3120
ctgacttaga aatagggcag catagaacaa aaatagagga actgagacaa catctgttga 3180
ggtggggatt taccacacca gacaaaaaac atcagaaaga acctccattc ctttggatgg 3240
gttatgaact ccatcctgat aaatggacag tacagcctat agtgctgcca gaaaaggaca 3300
gctggactgt caatgacata cagaaattag tgggaaaatt gaattgggca agtcagattt 3360
atgcagggat taaagtaagg caattatgta aacttcttag gggaaccaaa gcactaacag 3420
aagtagtacc actaacagaa gaagcagagc tagaactggc agaaaacagg gagattctaa 3480
aagaaccggt acatggagtg tattatgacc catcaaaaga cttaatagca gaaatacaga 3540
agcaggggca aggccaatgg acatatcaaa tttatcaaga gccatttaaa aatctgaaaa 3600
caggaaaata tgcaagaatg aagggtgccc acactaatga tgtgaaacaa ttaacagagg 3660
cagtacaaaa aatagccaca gaaagcatag taatatgggg aaagactcct aaatttaaat 3720
tacccataca aaaggaaaca tgggaagcat ggtggacaga gtattggcaa gccacctgga 3780
ttcctgagtg ggagtttgtc aatacccctc ccttagtgaa gttatggtac cagttagaga 3840
aagaacccat aataggagca gaaactttct atgtagatgg ggcagccaat agggaaacta 3900
aattaggaaa agcaggatat gtaactgaca gaggaagaca aaaagttgtc cccctaacgg 3960
acacaacaaa tcagaagact gagttacaag caattcatct agctttgcag gattcgggat 4020
tagaagtaaa catagtgaca gactcacaat atgcattggg aatcattcaa gcacaaccag 4080
ataagagtga atcagagtta gtcagtcaaa taatagagca gttaataaaa aaggaaaaag 4140
tctacctggc atgggtacca gcacacaaag gaattggagg aaatgaacaa gtagatgggt 4200
tggtcagtgc tggaatcagg aaagtactat ttttagatgg aatagataag gcccaagaag 4260
aacatgagaa atatcacagt aattggagag caatggctag tgattttaac ctaccacctg 4320
tagtagcaaa agaaatagta gccagctgtg ataaatgtca gctaaaaggg gaagccatgc 4380
atggacaagt agactgtagc ccaggaatat ggcagctaga ttgtacacat ttagaaggaa 4440
aagttatctt ggtagcagtt catgtagcca gtggatatat agaagcagaa gtaattccag 4500
cagagacagg gcaagaaaca gcatacttcc tcttaaaatt agcaggaaga tggccagtaa 4560
aaacagtaca tacagacaat ggcagcaatt tcaccagtac tacagttaag gccgcctgtt 4620
ggtgggcggg gatcaagcag gaatttggca ttccctacaa tccccaaagt caaggagtaa 4680
tagaatctat gaataaagaa ttaaagaaaa ttataggaca ggtaagagat caggctgaac 4740
atcttaagac agcagtacaa atggcagtat tcatccacaa ttttaaaaga aaaggggggga 4800
ttggggggta cagtgcaggg gaaagaatag tagacataat agcaacagac atacaaacta 4860
aagaattaca aaaacaaatt acaaaaattc aaaattttcg ggtttattac agggacagca 4920
gagatccagt ttggaaagga ccagcaaagc tcctctggaa aggtgaaggg gcagtagtaa 4980
tacaagataa tagtgacata aaagtagtgc caagaagaaa agcaaagatc atcagggatt 5040
atggaaaaca gatggcaggt gatgattgtg tggcaagtag acaggatgag gattaacaca 5100
tggaaaagat tagtaaaaca ccatatgtat atttcaagga aagctaagga ctggttttat 5160
agacatcact atgaaagtac taatccaaaa ataagttcag aagtacacat cccactaggg 5220
gatgctaaat tagtaataac aacatattgg ggtctgcata caggagaaag agactggcat 5280
ttgggtcagg gagtctccat agaatggagg aaaaagagat atagcacaca agtagaccct 5340
gacctagcag accaactaat tcatctgcac tattttgatt gtttttcaga atctgctata 5400
agaaatacca tattaggacg tatagttagt cctaggtgtg aatatcaagc aggacataac 5460
aaggtaggat ctctacagta cttggcacta gcagcattaa taaaaccaaa acagataaag 5520
ccacctttgc ctagtgttag gaaactgaca gaggacagat ggaacaagcc ccagaagacc 5580
aagggccaca gagggagcca tacaatgaat ggacactaga gcttttagag gaacttaaga 5640
gtgaagctgt tagacatttt cctaggatat ggctccataa cttaggacaa catatctatg 5700
aaacttacgg ggatacttgg gcaggagtgg aagccataat aagaattctg caacaactgc 5760
tgtttatcca tttcagaatt gggtgtcgac atagcagaat aggcgttact cgacagagga 5820
gagcaagaaa tggagccagt agatcctaga ctagagccct ggaagcatcc aggaagtcag 5880
cctaaaactg cttgtaccaa ttgctattgt aaaaagtgtt gctttcattg ccaagtttgt 5940
ttcatgacaa aagccttagg catctcctat ggcaggaaga agcggagaca gcgacgaaga 6000
gctcatcaga acagtcagac tcatcaagct tctctatcaa agcagtaagt agtacatgta 6060
atgcaaccta taatagtagc aatagtagca ttagtagtag caataataat agcaatagtt 6120
gtgtggtcca tagtaatcat agaatatagg aaaatattaa gacaaagaaa aatagacagg 6180
ttaattgata gactaataga aagagcagaa gacagtggca atgagagtga aggagaagta 6240
tcagcacttg tggagatggg ggtggaaatg gggcaccatg ctccttggga tattgatgat 6300
ctgtagtgct acagaaaaat tgtgggtcac agtctattat ggggtacctg tgtggaagga 6360
agcaaccacc actctatttt gtgcatcaga tgctaaagca tatgatacag aggtacataa 6420
tgtttgggcc acacatgcct gtgtacccac agaccccaac ccacaagaag tagtattggt 6480
aaatgtgaca gaaaatttta acatgtggaa aaatgacatg gtagaacaga tgcatgagga 6540
tataatcagt ttatgggatc aaagcctaaa gccatgtgta aaattaaccc cactctgtgt 6600
tagtttaaag tgcactgatt tgaagaatga tactaatacc aatagtagta gcgggagaat 6660
```

```
gataatggag aaaggagaga taaaaaactg ctctttcaat atcagcacaa gcataagaga 6720
taaggtgcag aaagaatatg cattcttta taaacttgat atagtaccaa tagataatac 6780
cagctatagg ttgataagtt gtaacacctc agtcattaca caggcctgtc caaaggtatc 6840
ctttgagcca attcccatac attattgtgc cccggctggt tttgcgattc taaaatgtaa 6900
taataagacg ttcaatggaa caggaccatg tacaaatgtc agcacagtac aatgtacaca 6960
tggaatcagg ccagtagtat caactcaact gctgttaaat ggcagtctag cagaagaaga 7020
tgtagtaatt agatctgcca atttcacaga caatgctaaa accataatag tacagctgaa 7080
cacatctgta gaaattaatt gtacaagacc caacaacaat acaagaaaaa gtatccgtat 7140
ccagagggga ccagggagag catttgttac aataggaaaa ataggaaata tgagacaagc 7200
acattgtaac attagtagag caaaatggaa tgccactta aaacagatag ctagcaaatt 7260
aagagaacaa tttggaaata ataaaacaat aatctttaag caatcctcag gaggggaccc 7320
agaaattgta acgcacagtt ttaattgtgg aggggaattt ttctactgta attcaacaca 7380
actgtttaat agtacttggt ttaatagtac ttggagtact gaagggtcaa ataacactga 7440
aggaagtgac acaatcacac tcccatgcag aataaaacaa tttataaaca tgtggcagga 7500
agtaggaaaa gcaatgtatg cccctcccat cagtggacaa attagatgtt catcaaatat 7560
tactgggctg ctattaacaa gagatggtgg taataacaac aatgggtccg agatcttcag 7620
acctggagga ggcgatatga gggacaattg gagaagtgaa ttatataaat ataaagtagt 7680
aaaaattgaa ccattaggag tagcacccac caaggcaaag agaagagtgg tgcagagaga 7740
aaaaagagca gtgggaatag gagctttgtt ccttgggttc ttgggagcag caggaagcac 7800
tatgggctgc acgtcaatga cgctgacggt acaggccaga caattattgt ctgatatagt 7860
gcagcagcag aacaatttgc tgagggctat tgaggcgcaa cagcatctgt tgcaactcac 7920
agtctggggc atcaaacagc tccaggcaag aatcctggct gtggaaagat acctaaagga 7980
tcaacagctc ctggggattt ggggttgctc tggaaaactc atttgcacca ctgctgtgcc 8040
ttggaatgct agttggagta ataaatctct ggaacagatt tggaataaca tgacctggat 8100
ggagtgggac agagaaatta caattacac aagcttaata cactccttaa ttgaagaatc 8160
gcaaaaccag caagaaaaga tgaacaaga attattggaa ttagataaat gggcaagttt 8220
gtggaattgg tttaacataa caaattggct gtggtatata aaattattca taatgatagt 8280
aggaggcttg gtaggtttaa gaatagtttt tgctgtactt tctatagtga atagagttag 8340
gcagggatat tcaccattat cgtttcagac ccacctccca atcccgaggg gacccgacag 8400
gcccgaagga atagaagaag aaggtggaga gagagacaga gacagatcca ttcgattagt 8460
gaacggatcc ttagcactta tctgggacga tctgcggagc ctgtgcctct tcagctacca 8520
ccgcttgaga gacttactct tgattgtaac gaggattgtg gaacttctgg gacgcagggg 8580
gtgggaagcc ctcaaatatt ggtggaatct cctacagtat tggagtcagg aactaaagaa 8640
tagtgctgtt aacttgctca atgccacagc catagcagta gctgaggcga cagatagggt 8700
tatagaagta ttacaagcag cttatagagc tattcgccac atacctagaa gaataagaca 8760
gggcttggaa aggattttgc tataagatgg gtggcaagtg gtcaaaaagt agtgtgattg 8820
gatggcctgc tgtaagggaa agaatgagac gagctgagcc agcagcagat ggggtgggag 8880
cagtatctcg agacctagaa aaacatggag caatcacaag tagcaataca gcagctaaca 8940
atgctgcttg tgcctggcta gaagcacaag aggaggaaga ggtgggtttt ccagtcacac 9000
ctcaggtacc tttaagacca atgacttaca aggcagctgt agatcttagc cacttttaa 9060
aagaaaaggg gggactggaa gggctaattc actcccaaag aagacaagat atccttgatc 9120
tgtggatcta ccacacacaa ggctacttcc ctgattggca gaactacaca ccagggccag 9180
gggtcagata tccactgacc tttggatggt gctacaagct agtaccagtt gagccagata 9240
aggtagaaga ggccaataaa ggagagaaca ccagcttgtt acaccctgtg agcctgcatg 9300
gaatggatga ccctgagaga gaagtgttag agtggaggtt tgacagccgc ctagcatttc 9360
atcacgtggc ccgagagctg catccggagt acttcaagaa ctgctgacat cgagcttgct 9420
acaagggact ttccgctggg gactttccag ggaggcgtgg cctgggcggg actggggagt 9480
ggcgagccct cagatgctgc atataagcag ctgcttttg cctgtactgg gtctctctgg 9540
ttagaccaga tctgagcctg ggagctctct ggctaactag ggaacccact gcttaagcct 9600
caataaagct tgccttgagt gcttcaagta gtgtgtgccc gtctgttgtg tgactctggt 9660
aactagagat ccctcagacc ctttttagtca gtgtggaaaa tctctagca 9709
```

<210> 2
<211> 854
<212> PRT
<213> Human immunodeficiency virus

<220>
<223> Envelope Polyprotein

<400> 2
Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Lys
 1   5   10   15

Trp Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Thr Glu
            20   25   30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
            35   40   45

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
        50   55   .   60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
65   70   75   80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp
                85   90   95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp
            100   105   110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser
            115   120   125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser
    130   135   140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn
145   150   155   160

Ile Ser Thr Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Phe Phe
                165   170   175

Tyr Lys Leu Asp Ile Val Pro Ile Asp Asn Thr Ser Tyr Arg Leu Ile
            180   -   185   190

Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe
    195   200   205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu
    210   215   220

Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val
225   230   235   240

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln
            245   250   255

Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Asp Val Val Ile Arg Ser
        260   265   270

Ala Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Thr
        275   280   285

```
Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser
    290                 295             300
Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys
305                 310             315                 320
Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp
            325             330             335
Asn Ala Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln Phe Gly
            340             345             350
Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu
        355             360             365
Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn
    370             375             380
Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr
385             390             395                 400
Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu Pro Cys
            405             410             415
Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val Gly Lys Ala Met
            420             425             430
Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr
        435             440             445
Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Gly Ser Glu
    450             455             460
Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu
465             470             475                 480
Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro
            485             490             495
Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly
        500             505             510
Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
        515             520             525
Gly Cys Thr Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
    530             535             540
Asp Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln
545             550             555                 560
Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
            565             570             575
Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
        580             585             590
Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp
        595             600             605
Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn Asn Met
    610             615             620
```

```
Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile
625             630             635                         640

His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln
            645             650                 655

Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asn
            660             665             670

Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile Val Gly
        675             680                 685

Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn
    690             695                 700

Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Leu Pro
705             710                 715                     720

Ile Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly
            725             730                 735

Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser Leu Ala
            740             745                 750

Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg
        755             760                 765

Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly
    770             775                 780

Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr
785             790                 795                     800

Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Asn Leu Leu Asn Ala Thr
            805             810                 815

Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val Leu Gln
            820             825                 830

Ala Ala Tyr Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg Gln Gly
        835             840                 845

Leu Glu Arg Ile Leu Leu
    850
```

<210> 3
<211> 107
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence:
      oligonucleotide for cloning

<400> 3
```
aagatgtagt aattagatct gccaatttca cagacaatgc taaaaccata atagtacagc  60
tgaacacatc gttagaaatt aattgtacaa gacccaacaa caataca              107
```

```
<210> 4
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence:
      oligonucleotide for cloning

<220>
<221> misc_feature
<222> (97)..(99)
<223> Sequence at this position: (GA)(AT)(GATC), ie.
      base at position 97 can be G or A, base at
      position 98 can be A or T, and base at
      position 99 can be G, A, T or C.

<400> 4
ttttgctcta gaaatgttac aatgtgcttg tcttatgtct cctgttgcag cttctgttgc   60
atgaaatgct ctccctggtc cgatatggat actatgrwnt tttcttgtat tgttgttggg  120


<210> 5
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence:
      sequencing primer

<400> 5
ccatgtacaa atgtcag                                                   17


<210> 6
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence:
      sequencing primer

<400> 6
aaaactgtgc gttacaa                                                   17


<210> 7
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence:
      sequencing primer

<400> 7
gtaaaacgac ggccagt                                                   17
```

```
<210> 8
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence:
      sequencing primer

<400> 8
caggaaacag ctatgac                                                    17


<210> 9
<211> 2148
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence: synthetic DNA

<220>
<221> misc_feature
<222> (3)..(9)
<223> BstEII cleavage site

<220>
<221> misc_feature
<222> (2143)..(2148)
<223> BamHI cleavage site

<400> 9
tgggtcaccg tctattatgg ggtgcctgtg tggaaggaag caaccaccac tctattttgt 60
gcatcagatg ctaaagcata tgatacagag gtacataatg tttgggccac acatgcctgt 120
gtacccacag accccaaccc acaagaagta gtattggtaa atgtgacaga aaattttaac 180
atgtggaaaa atgacatggt agaacagatg catgaggata taatcagttt atgggatcaa 240
agccttaagc catgtgtaaa attaacccca ctctgtgtta gtttaaagtg cactgatttg 300
aagaatgata ctaataccaa tagtagtagc gggagaatga taatggagaa aggagagata 360
aaaaactgca gcttcaatat cagcacaagc ataagagata aggtgcagaa agaatatgca 420
ttctttttata aacttgatat agtaccaata gataatacca gctataggtt gataagttgt 480
aacacctcag tgatcacaca ggcctgtcca aaggtatcct ttgagccaat tcccatacat 540
tattgtgccc cggctggttt tgcgattcta aatgtaata ataagacgtt caatggaaca 600
ggaccatgta caaatgtcag cacagtacaa tgtacacatg gaattcgacc agtagtatca 660
actcaactgc tgttaaatgg cagtctagca gaagaagatg tagtaattag atctgccaat 720
ttcacagaca atgctaaaac cataatagta cagctgaaca catctgtaga aattaattgt 780
acaagaccca acaacaatac aagaaaaagt atccgtatcc agaggggacc agggagagca 840
tttgttacaa taggaaaaat aggaaatatg agacaagcac attgtaacat ttctagagca 900
aaatggaatg ccacttttaaa acagatagct agcaaattaa gagaacaatt tggaaataat 960
aaaacaataa tctttaagca gtcatccgga ggggacccag aaattgtaac gcacagtttt 1020
aattgtggag gggaattttt ctactgtaat tcaacacaac tgtttaatag tacttggttt 1080
aatagtactt ggagtactga agggtcaaat aacactgaag gaagtgacac aatcacactc 1140
ccatgcagaa taaaacaatt tataaacatg tggcaggaag taggaaaagc aatgtatgcc 1200
cctcccatca gtggccaaat tagatgttca tcaaatatta ctgggctgct attaactcga 1260
gatggtggta ataacaacaa tgggtccgag attttcagac ctggaggagg cgatatgagg 1320
gataattgga gaagtgaatt atataaatat aaagtagtaa aaattgaacc attaggagta 1380
gcacccacca aggcaaagag acgcgtggtg cagagagaaa agcgcgcagt gggaatagga 1440
gctctgttcc ttgggttctt gggagcagca ggaagcacta tgggcgcagc gtcaatgacg 1500
ctgacggtac aggccagaca attattgtct gatatagtgc agcagcagaa caatttgctg 1560
agggcaattg aggcgcaaca gcatctgttg caactcacag tctggggcat caaacagctc 1620
caggcaagaa tcctggctgt ggaaagatac ctaaaggatc aacagctcct ggggatttgg 1680
ggttgctctg gaaaactcat ttgcaccact gctgtgcctt ggaatgctag ttggagtaat 1740
aaatctctgg aacagatttg gaataacatg acctggatgg agtgggacag agaaattaac 1800
aattacacaa gcttaataca ctccttaatt gaagaatcgc aaaaccagca agaaaagaat 1860
gaacaagaat tattggaatt agataaatgg gcaagtttgt ggaattggtt taacataaca 1920
aattggctgt ggtatataaa attattcata atgatagtag gaggcttggt aggtttaaga 1980
atagtttttg ctgtactttc tatagtgaat agagttaggc agggatattc accattatcg 2040
tttcagaccc acctcccaat cccgagggga cccgacaggc ccgaaggaat agaagaagaa 2100
```

ggtggagaga gagacagaga cagatccatt cgattagtga acggatcc 2148

```
<210> 10
<211> 6229
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of the artificial sequence: synthetic DNA

<220>
<221> sig_peptide
<222> (1293)..(1295)
<223> env ATG

<220>
<221> misc_feature
<222> (1377)..(1379)
<223> env AGT, gp120 start

<220>
<221> misc_feature
<222> (1397)..(1403)
<223> BstEII cleavage site

<220>
<221> misc_feature
<222> (3537)..(3542)
<223> BamHI cleavage site

<220>
<221> misc_feature
<222> (3855)..(3857)
<223> env TAA, stop

<400> 10
ctgacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga 60
ccgctacact tgccagcgcc ctagcgcccg ctcctttcgc tttcttccct tcctttctcg 120
ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctccctta gggttccgat 180
ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt tcacgtagtg 240
ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata 300
gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt 360
tataagggat tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt aacaaaaat 420
ttaacgcgaa ttttaacaaa atattaacgc ttacaatttc cattcgccat tcaggctgcg 480
caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 540
gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg 600
taaaacgacg gccagtgagc gtctagttat taatagtaat caattacggg gtcattagtt 660
catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga 720
ccgcccaacg accccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca 780
atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttggca 840
gtacatcaag tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg 900
cccgcctggc attatgccca gtacatgacc ttatgggact ttcctacttg gcagtacatc 960
tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacat caatgggcgt 1020
ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt 1080
ttgtttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg 1140
acgcaaatgg gcggtaggcg tgtacggtgg gaggtctata taagcagagc tcgtttagtg 1200
aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag aagacaccgg 1260
gacaattcga gctcggtacc gtcgacgcca ccatgagagt gaaggagaag tatcagcact 1320
tgtggagatg ggggtggaaa tggggcacca tgctccttgg gatattgatg atctgtagtg 1380
ctacagaaaa attgtgggtc accgtctatt atggggtacc tgtgtggaag gaagcaacca 1440
ccactctatt ttgtgcatca gatgctaaag catatgatac agaggtacat aatgtttggg 1500
ccacacatgc ctgtgtaccc acagacccca acccacaaga gtagtattg gtaaatgtga 1560
cagaaaattt taacatgtgg aaaaatgaca tggtagaaca gatgcatgag gatataatca 1620
gtttatggga tcaaagccta aagccatgtg taaaattaac cccactctgt gttagtttaa 1680
agtgcactga tttgaagaat gatactaata ccaatagtag tagcgggaga atgataatgg 1740
agaaaggaga gataaaaaac tgctctttca atatcagcac aagcataaga gataaggtgc 1800
```

```
agaaagaata tgcattcttt tataaacttg atatagtacc aatagataat accagctata 1860
ggttgataag ttgtaacacc tcagtcatta cacaggcctg tccaaaggta tcctttgagc 1920
caattcccat acattattgt gccccggctg gttttgcgat tctaaaatgt aataataaga 1980
cgttcaatgg aacaggacca tgtacaaatg tcagcacagt acaatgtaca catggaatca 2040
ggccagtagt atcaactcaa ctgctgttaa atggcagtct agcagaagaa gatgtagtaa 2100
ttagatctgc caatttcaca gacaatgcta aaaccataat agtacagctg aacacatctg 2160
tagaaattaa ttgtacaaga cccaacaaca atacaagaaa aagtatccgt atccagaggg 2220
gaccagggag agcatttgtt acaataggaa aaataggaaa tatgagacaa gcacattgta 2280
acattagtag agcaaaatgg aatgccactt taaaacagat agctagcaaa ttaagagaac 2340
aatttggaaa taataaaaca ataatcttta agcaatcctc aggaggggac ccagaaattg 2400
taacgcacag ttttaattgt ggaggggaat ttttctactg taattcaaca caactgttta 2460
atagtacttg gtttaatagt acttggagta ctgaagggtc aaataacact gaaggaagtg 2520
acacaatcac actcccatgc agaataaaac aatttataaa catgtggcag gaagtaggaa 2580
aagcaatgta tgcccctccc atcagtggac aaattagatg ttcatcaaat attactgggc 2640
tgctattaac aagagatggt ggtaataaca acaatgggtc cgagatcttc agacctggag 2700
gaggcgatat gagggacaat tggagaagtg aattatataa atataaagta gtaaaaattg 2760
aaccattagg agtagcaccc accaaggcaa agagaagagt ggtgcagaga gaaaaaagag 2820
cagtgggaat aggagctttg ttccttgggt tcttgggagc agcaggaagc actatgggct 2880
gcacgtcaat gacgctgacg gtacaggcca gacaattatt gtctgatata gtgcagcagc 2940
agaacaattt gctgagggct attgaggcgc aacagcatct gttgcaactc acagtctggg 3000
gcatcaaaca gctccaggca agaatcctgg ctgtggaaag atacctaaag gatcaacagc 3060
tcctggggat ttggggttgc tctggaaaac tcatttgcac cactgctgtg ccttggaatg 3120
ctagttggag taataaatct ctggaacaga tttggaataa catgacctgg atggagtggg 3180
acagagaaat taacaattac acaagcttaa tacactcctt aattgaagaa tcgcaaaacc 3240
agcaagaaaa gaatgaacaa gaattattgg aattagataa atgggcaagt ttgtggaatt 3300
ggtttaacat aacaaattgg ctgtggtata taaaattatt cataatgata gtaggaggct 3360
tggtaggttt aagaatagtt tttgctgtac tttctatagt gaatagagtt aggcagggat 3420
attcaccatt atcgtttcag acccacctcc caatcccgag gggacccgac aggcccgaag 3480
gaatagaaga agaaggtgga gagagagaca gagacagatc cattcgatta gtgaacggat 3540
ccttagcact tatctgggac gatctgcgga gcctgtgcct cttcagctac caccgcttga 3600
gagacttact cttgattgta acgaggattg tggaacttct gggacgcagg gggtgggaag 3660
ccctcaaata ttggtggaat ctcctacagt attggagtca ggaactaaag aatagtgctg 3720
ttaacttgct caatgccaca gccatagcag tagctgaggg gacagatagg gttatagaag 3780
tattacaagc agcttataga gctattcgcc acatacctag aagaataaga cagggcttgg 3840
aaaggatttt gctataagat gggtggcaag tggtcaaaaa gtagtgtgat tggatggcct 3900
gctgtaaggg aaagaatgag acgagctgag ccagcagcag atggggtggg agcagtatct 3960
cgagatctag actagaacta gcttcgatcc agacatgata agatacattg atgagtttgg 4020
acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat 4080
tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca 4140
ttttatgttt caggttcagg gggaggtgtg ggaggttttt taaagcaagt aaaacctcta 4200
caaatgtggt atggctgatt atgatcctgc ctcgcgcgtt tcggtgatga cggtgaaaac 4260
ctctgacaca tgcagctccc ggagacggtc acagcttgtc tgtaagcgga tgccgggagc 4320
agacaagccc gtcagggcgc gtcagcgggt gttggcgggt gtcggggcgc agccatgacc 4380
cagtcacgta gcgatagcgg agtgtatact ggcttaacta tgcggcatca gagcagattg 4440
tactgagagt gcaccatatg tcgggccgcg ttgctggcgt ttttccatag gctccgcccc 4500
cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta 4560
taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg 4620
ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc 4680
tcacgctgta ggtatctcag ttcggtgtag tcgttcgct ccaagctggg ctgtgtgcac 4740
gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac 4800
ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg 4860
aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga 4920
aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt 4980
agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag 5040
cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 5100
gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 5160
atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta agtatatat 5220
gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 5280
tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatatcc 5340
gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct 5400
ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca 5460
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 5520
ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg 5580
tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc 5640
cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag 5700
ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 5760
ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag 5820
```

```
  tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat 5880
  agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg 5940
  atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca 6000
  gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca 6060
  aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat 6120
  tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 6180
  aaaaataaac aaatagggggt tccgcgcaca tttccccgaa aagtgccac             6229
```

```
  <210> 11
  <211> 860
  <212> DNA
  <213> Human immunodeficiency virus

  <220>
  <221> misc_feature
  <222> (1)..(860)
  <223> PI-932 original sequence V1-V2-V3-loop

  <400> 11
  tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
  aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
  caaagcctaa agccatgtgt aaaactaacc ccactctgtg ttactttaaa ttgcactgat 180
  gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaat 240
  actcataaca atactattag tgggaaaata atagagaaag tagaaataaa aaactgctct 300
  ttcaaggtca ccacaggcat aagggataag atgcaaaaag aatatgcact tttgaataaa 360
  cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
  aacacctcga ccattacaca ggcctgtcca aaggtatcct ttgagccaat tcccatacat 480
  tttgtgccc cggctggttt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
  ggaccatgta aaaatgtcag cacagtacaa tgtacacatg gaattaggcc agtagtgtca 600
  actcaactgc tgttgaatgg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
  ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
  ataagacccc acaacactgt aacagacagg atacatatag gccagggag atcatttcat 780
  acaacaagaa aaataaaagg agatataaga caagcacatt gtagccttag gagaaaagat 840
  tggaataaca ctttacaaga                                            860
```

```
  <210> 12
  <211> 870
  <212> DNA
  <213> Artificial Sequence

  <220>
  <223> Description of the artificial sequence: PI-932
        gene cassette, comprising the cleavage sites for
        restriction enzymes BspTl, PstI, BclI, EcoRI,
        BglII, PvuII, XbaII, NheI

  <400> 12
  tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
  aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
  caaagcctta agccatgtgt aaaactaacc ccactctgtg ttactttaaa ttgcactgat 180
  gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaat 240
  actcataact gcagtattag tgggaaaata atagagaaag tagaaataaa aaactgctct 300
  ttcaaggtca ccacaggcat aagggataag atgcaaaaag aatatgcact tttgaataaa 360
  cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
  aacacctcgg tgatcacaca ggcctgtcca aaggtatcct ttgagccaat tcccatacat 480
  tttgtgccc cggctggttt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
  ggaccatgta aaaatgtcag cacagtacaa tgtacacatg gaattcggcc agtagtgtca 600
  actcaactgc tgttgaatgg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
  ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
  ataagacccc acaacactgt aacagacagg atacatatag gccagggag atcatttcat 780
  acaacaagaa aaataaaagg agatataaga caagcacatt gtagcctttc tagaaaagat 840
  tggaataaca ctttacaaga gatagctagc                                 870
```